7-2016

Robust median reversion strategy for online portfolio selection

Dingjiang HUANG

Junlong ZHOU

Bin LI

HOI, Steven C. H.
Singapore Management University, CHHOI@smu.edu.sg

Shuigeng ZHOU

DOI: https://doi.org/10.1109/TKDE.2016.2563433

Citation
HUANG, Dingjiang; ZHOU, Junlong; LI, Bin; HOI, Steven C. H.; and ZHOU, Shuigeng. Robust median reversion strategy for online portfolio selection. (2016). *IEEE Transactions on Knowledge and Data Engineering*. 28, (9), 2480-2493. Research Collection School Of Information Systems.
Available at: https://ink.library.smu.edu.sg/sis_research/3408

Published in IEEE Transactions on Knowledge and Data Engineering, 2016 Sept, Volume 28, Issue 9, pp. 2480-2493.
http://doi.org/10.1109/TKDE.2016.2563433

Robust Median Reversion Strategy for Online Portfolio Selection

Dingjiang Huang, Junlong Zhou, Bin Li, Steven C. H. Hoi, and Shuigeng Zhou

Abstract—Online portfolio selection has attracted increasing attention from data mining and machine learning communities in recent years. An important theory in financial markets is mean reversion, which plays a critical role in some state-of-the-art portfolio selection strategies. Although existing mean reversion strategies have been shown to achieve good empirical performance on certain datasets, they seldom carefully deal with noise and outliers in the data, leading to suboptimal portfolios, and consequently yielding poor performance in practice. In this paper, we propose to exploit the reversion phenomenon by using robust L_1-median estimators, and design a novel online portfolio selection strategy named "Robust Median Reversion" (RMR), which constructs optimal portfolios based on the improved reversion estimator. We examine the performance of the proposed algorithms on various real markets with extensive experiments. Empirical results show that RMR can overcome the drawbacks of existing mean reversion algorithms and achieve significantly better results. Finally, RMR runs in linear time, and thus is suitable for large-scale real-time algorithmic trading applications.

Index Terms—Portfolio selection, online learning, mean reversion, robust median reversion, L_1-median

1 Introduction

Portfolio Selection (PS) aims to determine an effective investment strategy for allocating wealth among a set of assets so as to achieve certain financial objectives in the long run. There are two main mathematical models for this task. The first is the mean-variance model [2], which trades off between a portfolio's expected return (mean) and risk (standard deviation), and is generally suitable for single-period (batch) PS. Another model is the Kelly investment (also termed "Capital Growth Theory") [3], which aims to maximize a portfolio's expected log return, and focuses mainly on multiple-period PS. These two theories have become the cornerstones of modern financial theory, whose principles are constantly visited and re-invented. One active research direction in data mining [4], [5], [6] and machine learning communities [7], [8] is online PS, which aims to design online algorithms following the Kelly model.

Some state-of-the-art online PS strategies [7], [9] assume that the current well performing stocks would continue to perform well in the next trading day, which is often known as the "momentum" principle. However, empirical evidence [10] indicates that such trends may be often violated, especially in the short term. This observation leads to the strategy of buying underperforming stocks and selling those over-performing ones, which is known as the "mean reversion" principle [11], [12].

Recent years have witnessed a surge of online PS studies [5], [8], [11], [12], [13] that have attempted to exploit the mean reversion principle. Although these algorithms achieve encouraging results on some datasets, they perform poorly on certain datasets, such as the DJA dataset [5], [8]. This is because real-world datasets often contain noisy data and outliers, while the existing mean reversion strategies do not fully address these issues, leading to estimation error and suboptimal portfolio (see [14]). Furthermore, the assumption of single-period prediction [5], [13] also leads to estimation errors and thus unsatisfactory performance [13].

To address the above drawbacks, we present a new multi-period online PS strategy named "Robust Median Reversion" (RMR). The basic idea is to exploit the reversion phenomenon via robust L_1-median estimators [15], [16], [17], which explicitly estimates the next price relative and is more accurate than traditional simple mean estimators. Then we learn optimal portfolios based on the improved reversion estimator and the state-of-the-art online learning techniques.

To the best of our knowledge, RMR is the first algorithm that exploits the reversion phenomenon by robust L_1-median estimator. Though simple in nature, it can achieve better estimates than the existing algorithms and has been empirically validated via extensive experiments on real markets. The experimental results show that RMR significantly surpasses a number of state-of-the-art strategies in terms of long-term compound return. Moreover, it is robust to different parameter settings and can withstand nontrivial transaction costs.

- *D.J. Huang is with the Department of Mathematics, East China University of Science and Technology, Shanghai 200237, China. E-mail: djhuang@ecust.edu.cn, djhuang.math@gmail.com.*
- *J.L. Zhou is with the Shanghai Futures Exchange, Shanghai 200122, China. E-mail: jlzhou8812@gmail.com.*
- *B. Li is with the Economics and Management School, Wuhan University Luojia Hill, Wuhan 430072 China. E-mail: binli.whu@whu.edu.cn.*
- *S.C.H. Hoi is with the School of Information Systems, Singapore Management University, 80 Stamford Road, Singapore 178902, Singapore. E-mail: chhoi@smu.edu.sg.*
- *S.G. Zhou is with the School of Computer Science and Shanghai Key Lab of Intelligent Information Processing, Fudan University, Shanghai 200433, China. E-mail: sgzhou@fudan.edu.cn.*

Finally, with a linear time complexity with respect to the number of stocks and the number of trading periods, RMR is suitable for large-scale applications.

As a summary, the main contributions of this paper include:

1) we propose a novel multi-period online PS strategy RMR to exploit the reversion phenomenon, which explicitly estimates the next price relative via robust L_1-median estimator and is more accurate than simple mean estimator;
2) we exploit two types of L_1-median estimators based on the absolute loss function and Huber loss function in order to deal with noise and outliers effectively and
3) we conduct extensive experiments to validate the performance of the proposed RMR algorithms by comparing with various state-of-the-art algorithms.

The rest of the paper is organized as follows. Section 2 formulates the online PS problem and Section 3 reviews some related work. Section 4 proposes the RMR algorithm and Section 5 empirically evaluates its efficacy on real markets. Section 6 finally summarizes this article.

2 Problem Setting

Now let us consider the online PS problem. We consider a financial market with d assets for n trading periods to be invested. On the tth period, the asset prices are represented by a *close price vector* $\mathbf{p}_t \in \mathbb{R}^d_+$, and each element p^i_t represents the close price of asset i. The changes of asset prices are represented by a *price relative vector* $\mathbf{x}_t = (x^1_t, \dots, x^d_t) \in \mathbb{R}^d_+$, where x^j_t expresses the ratio of close price to last close price of asset j at the tth period, i.e., $x^j_t = p^j_t / p^j_{t-1}$. We denote $\mathbf{x}^n_1 = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ as the sequence of price relative vectors for n periods.

At the beginning of the tth period, we allocate the capital among the d assets according to a *portfolio vector* $\mathbf{b}_t = (b^1_t, \dots, b^d_t)$, where b^j_t represents the proportion of wealth invested in the jth asset. Typically, we assume the portfolio is self-financed and no margin/short is allowed, which means each entry of a portfolio is non-negative and adds up to one, that is, $\mathbf{b}_t \in \Delta_d$, where $\Delta_d = \{\mathbf{b}_t : b^j_t \geq 0, \sum_{j=1}^d b^j_t = 1\}$. The investment procedure is represented by a *portfolio strategy*, that is, $\mathbf{b}_1 = \frac{1}{d}\mathbf{1}$ and the following sequence of mappings $\mathbf{b}_t : (\mathbb{R}^d_+)^{t-1} \to \Delta_d, t = 2, 3, \dots,$ where $\mathbf{b}_t = \mathbf{b}_t(\mathbf{x}^{t-1}_1)$ is the portfolio used on the tth trading period given past market sequence $\mathbf{x}^{t-1}_1 = (\mathbf{x}_1, \dots, \mathbf{x}_{t-1})$. We denote by $\mathbf{b}^n_1 = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ the strategy for n periods.

On the tth trading period, a portfolio $\mathbf{b}_t$ achieves a *portfolio period return* s_t, that is, the wealth increases by a factor of $s_t = \mathbf{b}^T_t \mathbf{x}_t = \sum_{j=1}^d b^j_t x^j_t$. Since we reinvest and adopt price relatives, the portfolio wealth would multiplicatively grow. Thus, after n trading periods, a portfolio strategy $\mathbf{b}^n_1$ produces a *portfolio cumulative wealth* S_n, which increases the initial wealth by a factor of $\prod_{t=1}^n \mathbf{b}^T_t \mathbf{x}_t$, that is, $S_n(\mathbf{b}^n_1, \mathbf{x}^n_1) = S_0 \prod_{t=1}^n (\mathbf{b}^T_t \mathbf{x}_t)$, where S_0 is the initial wealth, which is set to 1 in this paper.

Finally, we formulate the online PS problem as a sequential decision task following the aforementioned abstract problem. The portfolio manager aims to design a strategy $\mathbf{b}^n_1$ to maximize the portfolio cumulative wealth S_n. The portfolios are selected in a sequential fashion. On each period t, given the historical information, the manager learns to select a new portfolio vector $\mathbf{b}_t$ for the next price relative vector $\mathbf{x}_t$, where the decision criterion varies among different managers. The resulting portfolio $\mathbf{b}_t$ is scored by the portfolio period return of s_t. Such procedure repeats until the end of trading periods and the portfolio strategy is finally scored by the cumulative wealth S_n.

In the above model, we make several general assumptions:

1) Transaction cost: we assume no transaction cost or taxes in this PS model;
2) Market liquidity: we assume that one can buy and sell required quantities at last closing price of any given trading period;
3) Impact cost: we assume that market behavior is not affected by a PS strategy.

These assumptions are not trivial, which has been explained in many existing work (refer to Section 3 for detail). We will empirically analyze the effects of transaction costs in Section 5.4.

3 Related Work

In this section, we survey the online PS strategy literature from the point of estimated methods of the next price relatives. We start by introducing the benchmarks of the online PS task. After that, we categorize existing methods by their estimations of the next price relatives. Finally, we analyze existing algorithms following the framework of Kelly's investment. Readers are also encouraged to read a more comprehensive survey in [6], [18].

One common benchmark is *Buy And Hold* (BAH), which buys assets according to a pre-defined weight and holds until the end. In hindsight, the optimal BAH strategy over a sequence of price relatives is the *Best-stock*, which buys the best stock over the period. Another classical strategy is *Constantly Rebalanced Portfolios* (CRP), which keeps fixed weight on each asset for every period. In particular, the portfolio strategy can be represented as $\mathbf{b}^n_1 = (\mathbf{b}, \dots, \mathbf{b})$, in which $\mathbf{b}$ is a predefined portfolio. Note that CRP needs to rebalance the portfolio every period, while BAH does not. *Best CRP* (BCRP) [11], the best CRP strategy over a whole market sequence in hindsight, is an optimal strategy if the market is i.i.d. [19]. Thus, Cover [11] proposes to design a online PS strategy that asymptotically approaches the BCRP strategy.

The first category of existing methods is to estimate the next price relative via all historical price relatives with a uniform probability. This category includes *Successive Constantly Rebalanced Portfolios* (SCRP) [20] and *Online Newton Step* (ONS) [7].[1] Theoretically, SCRP has the same asymptotic growth of wealth as the BCRP and superior performance over portfolios which explicitly take into account possible nonstationary market behavior. ONS aims to maximize the expected logarithmic cumulative wealth (approximated using historical price relatives) and minimize the variation of the expected portfolio. Since it iteratively updates the first and second order information, it costs

1. SCRP and ONS's formulations are similar, while they use different techniques to solve the formulations.

$O(d^3)$ per period, which is irrelevant to the number of past periods. Here d denotes the number of stocks.

The second category of strategies predicts the next price relatives via a set of similar historical price relatives. These strategies contain a pattern matching step, which selects the set of similar price relatives, and a portfolio optimization step, which constructs an optimal portfolio. *Nonparametric kernel based moving window* (B^K) [21] compares the patterns using Euclidean distance, and constructs an optimal portfolio as the BCRP on the obtained set of price relatives. Following the same framework, *Nonparametric Nearest Neighbor* (B^{NN}) [22] locates the set of price relatives via nearest neighbor methods, and *Correlation-driven Nonparametric learning* (CORN) [23] measures the similarity via correlation. Recently these results have been summarized and extended by Györfi et al. in their published book [24].

Moreover, the third category of estimation is to predict the next price relative via a single-value prediction. *Exponential Gradient* (EG) [9] estimates the next price relative as the last price relative. *Passive Aggressive Mean Reversion* (PAMR) [13] and *Confidence Weighted Mean Reversion* (CWMR) [5] estimate next price as the inverse of last price relative, which is in essence the "mean reversion" principle[2]. Recently, [8] proposed *Online Moving Average Reversion* (OLMAR), which predicts the next price relative using moving averages and explores the multi-period mean reversion.

Finally, some algorithms do not focus on estimation, either explicitly or implicitly. *Universal portfolios* (UP) [11] is the historical performance weighted average of all CRPs. *Anti-Correlation* (Anticor) [12] adopts the consistency of positive lagged cross-correlation and negative autocorrelation to adjust the portfolio. There are also some algorithms which focus on transaction cost. *Online lazy updates* (OLU) [25] and *online lazy updates with group sparsity* (OLU-GS) [26] with transaction cost take advantage of EG algorithm [9], which rebalances the portfolio vector by *lazy* or *sparse* updates of the parameters in the optimization model. *Semi-Constant rebalanced portfolio* (SCRP) [27] and *Semi-Universal portfolios* (SUP) [28] with transaction cost combine CRP and UP with occasional trading, respectively.

3.1 Analysis of Existing Work

Now, let us focus on the estimation methods of existing work. In practice, a Kelly portfolio manager [3], [29] firstly predicts $\hat{\mathbf{x}}_{t+1}$ in terms of k possible values $\hat{\mathbf{x}}^1_{t+1}, \ldots, \hat{\mathbf{x}}^k_{t+1}$ and their corresponding probabilities $p_1, \ldots, p_k$, where each $\hat{\mathbf{x}}^i_{t+1}$ denotes one possible combination vector of individual price relative predictions. Then he/she can figure out a portfolio by maximizing the expected log return on the possible combinations,

$$\mathbf{b}_{t+1} = \underset{\mathbf{b}\in\Delta_d}{\arg\max} \sum_{i=1}^{k} p_i \log\left(\mathbf{b}^T \hat{\mathbf{x}}^i_{t+1}\right).$$

As different estimation methods give different $\hat{\mathbf{x}}^i_{t+1}$ and p_i and lead to different portfolios, an accurate estimation method is crucial to the success of a strategy.

2. PAMR and CWMR adopt the same estimation, while they exploit the principle via different techniques.

Below, we analyze mainly the algorithms PAMR, CWMR and OLMAR, which estimate the next price relative by a single value prediction based on mean reversion or moving average reversion. PAMR and CWMR implicitly assume $\hat{\mathbf{x}}^1_{t+1} = \frac{1}{\mathbf{x}_t}$ with $p_1 = 100\%$ i.e., they estimate the next price relative as the inverse of last price relative, which is in essence the mean reversion principle. From the price perspective [13], they implicitly assume that next price $\hat{\mathbf{p}}_{t+1}$ will revert to last price $\mathbf{p}_{t-1}$,

$$\hat{\mathbf{x}}_{t+1} = \frac{1}{\mathbf{x}_t} \Rightarrow \frac{\hat{\mathbf{p}}_{t+1}}{\mathbf{p}_t} = \frac{\mathbf{p}_{t-1}}{\mathbf{p}_t} \Rightarrow \hat{\mathbf{p}}_{t+1} = \mathbf{p}_{t-1},$$

where $\mathbf{x}$ and $\mathbf{p}$ are all vectors and the above operations are element-wise. Rather than $\hat{\mathbf{p}}_{t+1} = \mathbf{p}_{t-1}$, OLMAR estimates the next price as a moving average at the end of the tth period, that is, $\hat{\mathbf{p}}_{t+1} = MA_t(w) = \frac{1}{w}\sum_{i=t-w+1}^{i=t} \mathbf{p}_i$ where $MA_t(w)$ denotes the moving average with a w-window. Though these estimation methods in PAMR/CWMR and OLMAR are empirically effective on most datasets, they do have potential problems. Firstly, PAMR and CWMR adopt the single-period mean reversion assumption in designing the algorithms, which is not always satisfied in the real world. One real example [13] is the DJA dataset, on which PAMR performs the worst among the state of the art. Secondly, all three algorithms suffer from the frequently fluctuating raw prices, which often contain a lot of noise and outliers and thus substantially influences the effectiveness of the algorithm and even the final cumulative wealth. Considering these drawbacks of the existing works, we try to develop a new robust reversion strategy.

It should be noted that in this paper we don't consider the transaction costs in our original algorithmic formulation. Generally, there are two ways to deal with transaction cost in the online PS. The first way is that one does not consider the transaction cost during the PS process and evaluate the impact of transaction costs in the back tests. This way has been commonly adopted in the designing a online PS strategies [8], [12], [30]. The second way is that the transaction cost is directly involved in the PS process [31]. The online PS strategies related to this way is usually named transaction cost aware strategies, such as OLU, SCRP and SUP [25], [26], [27], [28], and so on. Therefore, the strategies proposed in this paper belongs to the first way and thus we did not compare it with those transaction costs aware ones in the later experiments. More details can be found in the Section 5.2.3.

4 Robust Median Reversion

In this section, we first give a motivating example and then present the proposed RMR and RMR-Variant strategies.

4.1 Motivation

Empirical evidence [5], [13] shows that if the asset price follows a normal distribution, the mean of historical prices is the statistical optimal estimate. OLMAR, which estimates the next price via a moving average, also achieves good results on most datasets. However, due to noise and outliers in the real market data [14] and the real markets are not

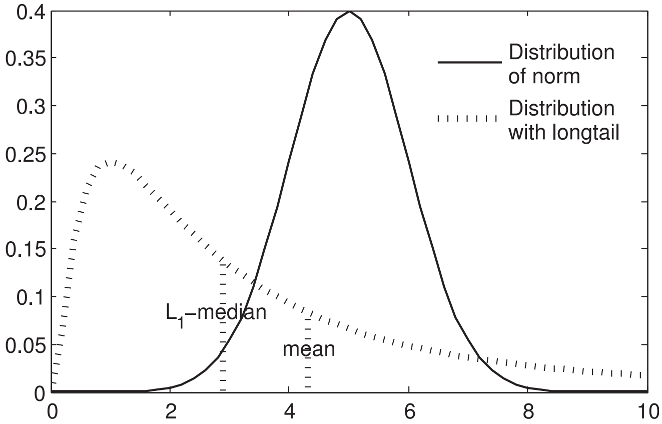

Fig. 1. Normal distribution and distribution with longtail.

TABLE 1
Illustration of Different Price Estimation Methods on Toy Markets

Price:$t_0 \rightarrow t_1 \rightarrow \ldots$	Acc	PAMR/CWMR	OLMAR	RMR
$A_0: 1, 2, 1, 2, ?$	1	1	1.5	1.5
$A_1: 1, 2, (10), 2, ?$	1	10	3.75	2
$B_0: 1, 2, 4, 2, ?$	1	4	2.25	2
$B_1: 1, 2, (10), 2, ?$	1	10	3.75	2
$C_0: 1, 2, 4, 8, 4, 2, ?$	1	4	3.5	3
$C_1: 1, 2, 4, 8, (10), 2, ?$	1	10	4.5	3

A_0, A_1; B_0, B_1 and C_0, C_1 represent single-period, two-period and three-period price sequence, respectively. A_0, B_0, C_0 are exact price sequence, and A_1, B_1, C_1 are price sequence contaminated by an outlier of 10. "Acc" is the accurate price. Other three items represent three estimates based on three different methods.

normally distributed, the price distribution often has a long tail[3] (see Fig. 1), and the previous estimation methods are sub-optimal subject to the noise and outliers.

To illustrate the drawbacks of mean and moving average, let us see a toy example. The toy market consists of one volatile stock, and $t_i (i \geq 0)$ denotes the period that requires estimation. Several sequences of market prices are listed in Table 1. A_0 and A_1 are single-period price sequences and their prices change by sequent factors of $2, \frac{1}{2}, 2, \frac{1}{2}, \ldots$ For example, let P_{t_i} be the price of the *i*th period, then $P_{t_1}{=}P_{t_0}{\times}2{=}1{\times}2{=}2, P_{t_2}{=}P_{t_1}{\times}\frac{1}{2}{=}2{\times}\frac{1}{2}{=}1, P_{t_3}{=}P_{t_2}{\times}2 = 1{\times}2{=}2, \ldots$ B_0 and B_1 are two-period price sequences and their prices change by sequent factor of $2, 2, \frac{1}{2}, \frac{1}{2}, 2, 2, \ldots C_0$ and C_1 are the three-period price sequences, and the price changes by sequent factor of $2, 2, 2, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 2, 2, 2, \ldots$ Moreover, A_0, B_0 and C_0 are exact price sequences, while A_1, B_1 and C_1 are the sequences contaminated by an outlier of (10). "?" denotes the price to be estimated and "Acc" is the accurate price. The estimated prices clearly show that the next prices estimated by PAMR/CWMA and OLMAR are far away from the accurate values, which thus leads to inaccurate price relatives and sub-optimal portfolios.

To better exploit the (multiple period) reversion property, we propose two new types of estimation methods and a new type of algorithm for online PS, named "*Robust Median Reversion*" . The essential idea is to exploit multiple period reversion via robust L_1-median estimator [15], [16], [17] and online machine learning. Rather than $\hat{\mathbf{p}}_{t+1} = \mathbf{p}_{t-1}$ or $\hat{\mathbf{p}}_{t+1} = MA_t(w)$, RMR estimates the next price by robust L_1-median estimator at the end of t^{th} period, that is, $\hat{\mathbf{p}}_{t+1} = L_1 med_{t+1}(w) = \boldsymbol{\mu}$, where w is the window size, $\boldsymbol{\mu}$ denotes the L_1-median estimator optimal value of Optimization Problems 1 and 2 below (see Section 4.2). Then the expected price relative with L_1-median estimator is

$$\hat{\mathbf{x}}_{t+1}(w) = \frac{L_1 med_{t+1}(w)}{\mathbf{p}_t} = \frac{\boldsymbol{\mu}}{\mathbf{p}_t}. \tag{1}$$

Without detailing the calculation,[4] we list the estimated next price of RMR in different toy markets in Table 1. Clearly, for the multiple period price sequences B_0, B_1 and C_0, C_1, RMR estimate is much closer to the Accurate values than PAMR/CWMR, showing that RMR method can deal with multiple period price sequence. For the contaminated price sequences A_1, B_1, C_1, RMR is also closer to the Accurate values than OLMAR and PAMR/CWMR estimates which implies that RMR is a robust method. Hence, the proposed methods provide better estimates and subsequent better portfolios than mean and moving average estimations. Note that although the toy example is on a single asset, such estimate goodness can be easily extended to the scenario of multiple assets.

4.2 Formulation

The proposed formulation, RMR, is to exploit median reversion via robust L_1-estimator and Passive Aggressive online learning [33]. The basic idea is to obtain the next price relative $\hat{\mathbf{x}}_{t+1}$ using multivariate L_1-median, and then maximize the expected return $\mathbf{b} \cdot \hat{\mathbf{x}}_{t+1}$ while keeping last portfolio information via regularization.

In statistics, the L_1-median (also named spatial median) [34], [35] is solution of the problem of minimizing the weighted sum of the Euclidean distances from k points in $\mathbb{R}^n$. This problem can be formulated in an even more general form by Weber [15] (Fermat-Weber problem), referred as location issues in industrial applications. In this article, L_1-median is the point with minimal sum of Euclidean distances to the k given price data points. To calculate the multivariate L_1-median of a k-historical price window, we adopt two types of L_1-median estimator. The first is so called "Minimum-Average Absolute Deviation Median" (MAADM), which satisfy the following optimization:

Optimization Problem 1: L_1-Median-MAADM

$$\boldsymbol{\mu} = \underset{\boldsymbol{\mu}}{\operatorname{argmin}} \sum_{i=0}^{k-1} \left\| \mathbf{p}_{t-i} - \boldsymbol{\mu} \right\|, \tag{2}$$

where $\|\cdot\|$ denotes the Euclidean norm. This L_1-median estimator is one of the classical statistics, which are robust to outliers and noisy data. The minimum-average absolute deviation median in Optimization problem 1 minimizes the loss with respect to the absolute loss function.

There also exist some other loss functions which are related to robust statistics [36]. One of such functions is

3. Fama have also indicated that the stock price is a heavy tail in his well-known paper [32].

4. We calculate RMR's expected price relative using Algorithm 1.

Huber loss function, which is used to construct an estimate that reduces the effect of outliers, while treating non-outliers in a more standard way. Thus, to reduce the effect of outliers and noisy data, we consider the second type of robust median estimator, named "Huber Loss Function-Based Median" (HLFBM), which replace the absolute loss function in Eq. (2) with the Huber loss function [37]:

Optimization Problem 2 : L_1-Median-HLFBM

$$\boldsymbol{\mu} = \underset{\boldsymbol{\mu}}{\operatorname{argmin}} \sum_{i=0}^{k-1} \rho\left(\|\mathbf{p}_{t-i} - \boldsymbol{\mu}\|\right), \tag{3}$$

where $\|\cdot\|$ denotes the Euclidean norm. Here, $\rho()$ is the Huber loss function defined as:

$$\rho(\gamma) = \begin{cases} \gamma^2/2, & |\gamma| \leq c \\ c(|\gamma| - c/2), & |\gamma| > c \end{cases}. \tag{4}$$

The Huber loss function is quadratic for small values of γ, and linear for large values, with equal values when $|\gamma| = c$ [36], [37].

To this end, we can calculate the expected price relative following the idea of so called "Median Reversion" (MR). Based on the two types of median estimators, we can infer two types of MR by Eq. (1),

Median Reversion: MR

$$\hat{\mathbf{x}}_{t+1}(w) = \frac{L_1 medianMAADM_{t+1}(w)}{\mathbf{p}_t} = \frac{\boldsymbol{\mu}}{\mathbf{p}_t}, \tag{5}$$

where w is the window size, $\boldsymbol{\mu}$ denotes the value in L_1-median estimator that satisfied the Optimization Problem 1.

Median Reversion: MR-Variant

$$\hat{\mathbf{x}}_{t+1}(w) = \frac{L_1 medianHLFBM_{t+1}(w)}{\mathbf{p}_t} = \frac{\boldsymbol{\mu}}{\mathbf{p}_t}, \tag{6}$$

where w is the window size, $\boldsymbol{\mu}$ denotes the value in L_1-median estimator that satisfied the Optimization Problem 2.

Based on the obtained price relative $\hat{\mathbf{x}}_{t+1}$ in (5) and (6), RMR further adopts the idea of an effective online learning algorithm, that is, Passive Aggressive (PA) learning [33], to exploit median reversion. Generally proposed for classification, PA passively keeps the previous solutions if the classification is correct, while aggressively approaches a new solution if the classification is incorrect. After formulating the proposed RMR optimization, we solve its closed form update and design the proposed algorithm.

The proposed formulation, RMR, is to exploit median reversion via online learning techniques. The basic idea is to maximize the expected return $\mathbf{b} \cdot \hat{\mathbf{x}}_{t+1}$ while keeping last portfolio information via regularization using the online passive-aggressive method. Thus, following the similar idea PAMR and OLMAR [8], [13], we can formulate the following optimization,

Optimization Problem 3 : RMR

$$\mathbf{b}_{t+1} = \underset{\mathbf{b} \in \Delta_d}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{b} - \mathbf{b}_t\|^2 \quad s.t. \ \ \mathbf{b}^T \hat{\mathbf{x}}_{t+1} \geq \varepsilon. \tag{7}$$

The above formulation attempts to find an optimal portfolio by minimizing the deviation from last portfolio $\mathbf{b}_t$ under the condition of $\mathbf{b}^T \hat{\mathbf{x}}_{t+1} \geq \varepsilon$. Such formulation explicitly reflects the reversion idea underlying the proposed RMR. In fact, $\hat{\mathbf{x}}_{t+1}$ is the price relative estimated via L_1-median estimators, while the constraint $\mathbf{b}^T \hat{\mathbf{x}}_{t+1} \geq \varepsilon$ means that next price will revert to the L_1-median. On the one hand, if its constraint is satisfied, that is, the expected return is higher than a threshold, then the resulting portfolio equals to previous portfolio. On the other hand, if the constraint is not satisfied, then the formulation will figure out a new portfolio such that the expected return is higher than the threshold, while the new portfolio is not far from previous one.

Remark on median estimators. Note the L_1-median estimator is much better than mean estimators statistically. In fact, the L_1-median has an attractive statistical properties, that is, its breakdown point is 0.5 [38], i.e., only if more than 50 percent of the data points are contaminated, the L_1-median can take values beyond all bounds. Note that breakdown point, the proportion of incorrect observations an estimator can handle, is a statistical metric of robustness. The higher the breakdown point of an estimator is, the more robust it is. However, the breakdown point of mean is 0, which means that the mean estimator is sensitive to the noisy data and outliers.

4.3 Algorithm

To obtain the L_1-median-MAADM of historical prices, we solve its optimization via the Modified Weiszfeld Algorithm [17], which converges monotonically to the L_1-median-MAADM. The solution of L_1-median-MAADM described in Eq. (2) is illustrated in Proposition 1. Its derivations are included in the Appendix.

Proposition 1. *The solution of L_1-median-MAADM optimization problem 1 is calculated through iteration, and the iteration process is described as:*

$$\boldsymbol{\mu} \to T(\boldsymbol{\mu}) = \left(1 - \frac{\eta(\boldsymbol{\mu})}{\gamma(\boldsymbol{\mu})}\right)^+ \tilde{T}(\boldsymbol{\mu}) + \min\left(1, \frac{\eta(\boldsymbol{\mu})}{\gamma(\boldsymbol{\mu})}\right)\boldsymbol{\mu},$$

where

$$\eta(\boldsymbol{\mu}) = \begin{cases} 1 & if \ \boldsymbol{\mu} = \mathbf{p}_{t-i}, \quad i = 0, \ldots, k-1 \\ 0 & otherwise \end{cases},$$
$$\gamma(\boldsymbol{\mu}) = \|\tilde{R}(\boldsymbol{\mu})\|, \quad \tilde{R}(\boldsymbol{\mu}) = \sum_{\mathbf{p}_{t-i} \neq \boldsymbol{\mu}} \frac{\mathbf{p}_{t-i} - \boldsymbol{\mu}}{\|\mathbf{p}_{t-i} - \boldsymbol{\mu}\|},$$
$$\tilde{T}(\boldsymbol{\mu}) = \left\{\sum_{\mathbf{p}_{t-i} \neq \boldsymbol{\mu}} \frac{1}{\|\mathbf{p}_{t-i} - \boldsymbol{\mu}\|}\right\}^{-1} \sum_{\mathbf{p}_{t-i} \neq \boldsymbol{\mu}} \frac{\mathbf{p}_{t-i}}{\|\mathbf{p}_{t-i} - \boldsymbol{\mu}\|}.$$

In general, we often practically set the convergence criterion during the iteration. Here, once the constraint of $\|\boldsymbol{\mu}_{t-1} - \boldsymbol{\mu}_t\|_1 \leq \tau \|\boldsymbol{\mu}_t\|_1$ is satisfied, we terminate the iteration. Note that τ represents a toleration level.

To obtain the L_1-median-HLFBM of historical prices, we apply the Quasi-Newton Algorithm [39]. In this algorithm, we update the Hessian matrix by BFGS [40], and calculate the scalar step length parameter by polynomial interpolation line search based on Wolfe principle. The solution of L_1-median-HLFBM described in Eq. (3) is illustrated below. We omit the detailed derivations.

Proposition 2. *The solution of L_1-median-HLFBM Optimization problem 2 is calculated through iteration, and the iteration process is described as:*

$$\boldsymbol{\mu}_{i+1} = \boldsymbol{\mu}_i + \alpha_i \boldsymbol{d}_i$$

where

$$\boldsymbol{d}_i = -B_i^{-1} \bigtriangledown f(\boldsymbol{\mu}_i),$$

$$B_i = TB(B_{i-1}) = B_{i-1} - \frac{B_{i-1}\boldsymbol{s}_{i-1}\boldsymbol{s}_{i-1}^T B_{i-1}}{\boldsymbol{s}_{i-1}^T B_{i-1}\boldsymbol{s}_{i-1}} + \frac{\boldsymbol{y}_{i-1}\boldsymbol{y}_{i-1}^T}{\boldsymbol{s}_{i-1}^T \boldsymbol{y}_{i-1}}$$

$$\boldsymbol{s}_{i-1} = \boldsymbol{\mu}_i - \boldsymbol{\mu}_{i-1}, \quad \boldsymbol{y}_{i-1} = \nabla f(\boldsymbol{\mu}_i) - \nabla f(\boldsymbol{\mu}_{i-1}),$$

and α_i is calculated by polynomial interpolation line search algorithm.

In general, we often practically set the convergence criterion during the iteration. Here, once the constraint of $\bigtriangledown f(\boldsymbol{\mu}_i) \leq \tau$ is satisfied, we terminate the iteration. Note that τ represents convergence precision.

After obtaining the next price relative, we can obtain the final PS formula by solving the Optimization problem 3, which is convex and thus straightforward to solve via the Lagrange multiplier method [41]. Its derivations are included in the Appendix.

Algorithm 1. L_1-Median-MAADM$(\mathbf{p}_t, \ldots, \mathbf{p}_{t-w+1}, m, \tau)$

Input: data $\mathbf{p}_t, \mathbf{p}_{t-1}, \ldots, \mathbf{p}_{t-w+1}$; iteration maximum m; toleration level τ
Output: estimated $\hat{\mathbf{x}}_{t+1}$
Procedure:
Initialize $\boldsymbol{\mu}_1 = median(\mathbf{p}_t, \mathbf{p}_{t-1}, \ldots, \mathbf{p}_{t-w+1})$.
for $i = 2$ **to** m **do**
  $\boldsymbol{\mu}_i = T(\boldsymbol{\mu}_{i-1})$
  if $\|\boldsymbol{\mu}_{i-1} - \boldsymbol{\mu}_i\|_1 \leq \tau \|\boldsymbol{\mu}_i\|_1$ **then**
    break
  end if
end for
$\hat{\mathbf{p}}_{t+1} = \boldsymbol{\mu}_i$
$\hat{\mathbf{x}}_{t+1} = \hat{\mathbf{p}}_{t+1} / \mathbf{p}_t$

Algorithm 2. L_1-Median-HLFBM$(\mathbf{p}_t, \ldots, \mathbf{p}_{t-w+1}, m, \tau)$

Input: data $\mathbf{p}_t, \mathbf{p}_{t-1}, \ldots, \mathbf{p}_{t-w+1}$; iteration maximum m; toleration level τ
Output: estimated $\hat{\mathbf{x}}_{t+1}$
Procedure:
Initialize $\boldsymbol{\mu}_1 = I_d$. $B_1 = I_{d\times d}$.
while $\bigtriangledown f(\boldsymbol{\mu}_i) > \tau$ and $i < m$ **do**
  $\boldsymbol{d}_i = -B_i^{-1} \bigtriangledown f(\boldsymbol{\mu}_i)$
  $\alpha_i =$ value by polynomial interpolation line search
  $\boldsymbol{u}_{i+1} = \boldsymbol{\mu}_i + \alpha_i \boldsymbol{d}_i$
  $B_{i+1} = TB(B_i)$
  $i = i + 1$
end while
$\hat{\mathbf{p}}_{t+1} = \boldsymbol{\mu}_i$
$\hat{\mathbf{x}}_{t+1} = \hat{\mathbf{p}}_{t+1} / \mathbf{p}_t$

Proposition 3. *The solution of RMR Optimization problem 3 without considering the non-negativity constraint is*

$$\mathbf{b}_{t+1} = \mathbf{b}_t - \alpha_{t+1}(\hat{\mathbf{x}}_{t+1} - \overline{x}_{t+1} \cdot \mathbf{1}),$$

where $\overline{x}_{t+1} = \frac{1}{d}(\mathbf{1} \cdot \hat{\mathbf{x}}_{t+1})$ denotes the average predicted price relative and α_{t+1} is the Lagrangian multiplier calculated as,

$$\alpha_{t+1} = \min\left\{0, \frac{\hat{\mathbf{x}}_{t+1}^T \mathbf{b}_t - \epsilon}{\|\hat{\mathbf{x}}_{t+1} - \overline{x}_{t+1} \cdot \mathbf{1}\|^2}\right\}.$$

Note that it is possible that the obtained portfolio in Proposition 3 goes out of the simplex domain since we do not consider the non-negativity constraint following [9]. Thus, to ensure that the portfolio is non-negative, we finally project the above portfolio to the simplex domain [42].

Algorithm 3. RMR$(\epsilon, \hat{\mathbf{x}}_{t+1}, \mathbf{b}_t)$

Input: reversion threshold $\epsilon > 1$; predicted the next price relative vector $\hat{\mathbf{x}}_{t+1}$; current portfolio $\mathbf{b}_t$;
Output: next portfolio $\mathbf{b}_{t+1}$
Procedure:
Calculate the following variable:

$$\alpha_{t+1} = \min\left\{0, \frac{\hat{\mathbf{x}}_{t+1}^T \mathbf{b}_t - \epsilon}{\|\hat{\mathbf{x}}_{t+1} - \overline{x}_{t+1} \cdot \mathbf{1}\|^2}\right\}$$

Update the portfolio:

$$\mathbf{b}_{t+1} = \mathbf{b}_t - \alpha_{t+1}(\hat{\mathbf{x}}_{t+1} - \overline{x}_{t+1} \cdot \mathbf{1})$$

Normalize $\mathbf{b}_{t+1}$: $\mathbf{b}_{t+1} = argmin_{\mathbf{b}\in\Delta_d} \|\mathbf{b} - \mathbf{b}_{t+1}\|^2$

Algorithm 4. PS with RMR and RMR-Variant

Input: reversion threshold $\epsilon > 1$; iteration maximum m; window size $w \geq 2$; toleration level τ; market sequence $\mathbf{x}_1^n$
Output: S_n: Cumulative wealth after nth periods
Procedure:
Initialization: $b_1 = \frac{1}{d}\mathbf{1}, S_0 = 1, \mathbf{p}_0 = \mathbf{1}$
for $t = 1, 2, \ldots, n$ **do**
  Receive stock price: $\mathbf{x}_t$
  Update cumulative return: $S_t = S_{t-1} \times (\mathbf{b}_t \cdot \mathbf{x}_t)$
  Predict the next price relative vector:

$$\hat{\mathbf{x}}_{t+1} = \begin{cases} \frac{L_1 medianMAADM_{t+1}(w)}{\mathbf{p}_t} & \text{MR} \\ \frac{L_1 medianHLFBM_{t+1}(w)}{\mathbf{p}_t} & \text{MR-Variant} \end{cases}$$

  Update the portfolio:

$$\mathbf{b}_{t+1} = \mathrm{R}MR(\epsilon, \hat{\mathbf{x}}_{t+1}, \mathbf{b}_t)$$

end for

To this end, we can design the proposed algorithms based on the above Propositions. The estimated process of price relative $\hat{\mathbf{x}}_{t+1}$, mainly based on Proposition 1, is illustrated in Algorithm 1. The estimated process of price relative $\hat{\mathbf{x}}_{t+1}$, mainly based on Proposition 2, is illustrated in Algorithm 2. The proposed RMR procedure, following Proposition 3, is shown in Algorithm 3. Finally, Algorithm 4 presents the

TABLE 2
Summary of Time Complexity Analysis

Methods	Time Complexity	Methods	Time Complexity
UP	$O(n^d)/O(d^7n^8)$	ONS	$O(d^3n)$
EG	$O(dn)$	Anticor	$O(N^3d^2n)$
PAMR/CWMR /OLMAR	$O(dn)$	$\mathrm{B^K}/\mathrm{B^{NN}}$ /CORN	$O(N^2dn^2)$ $+O(Ndn^2)$
RMR	$O(dn)+O(mn)$		

d denotes the number of stocks; n is the number of trading periods; N denotes the number of experts; and m denotes the number of loops in Algorithm 1.

online PS RMR and RMR-Variant which correspond to the two types median reversion MR and MR-Variant strategies.

4.4 Complexity Analysis

It is widely known that computational time is important to certain trading environments, such as high frequency trading [43], where trades occur in fractions of a second. RMR's time complexity is linear with respect to d and n, where n is much larger than d. In the RMR implementation, the max number of loop (Line 6 in Algorithm 1) can be implemented in $O(m)$. Thus, Algorithm 1 take $O(m)$ time per period. Moreover, Algorithm 3 takes $O(d)$ per period. In total, the whole time complexity is $O(dn)+O(mn)$. Table 2 compares the computational time complexity of RMR with that of existing strategies. Clearly, the proposed RMR algorithm takes no more time than any others.

Remark on theoretical analysis. It is true that we only provide empirical results for RMR and RMR-Variant in this paper. RMR and RMR-Variant adopt the median reversion property, which distinguishes the algorithms with the theoretical guaranteed algorithms, such as UP/EG/ONS. The property leads to the excellent empirical performance, however, prevents us from providing the theoretical results. Such claims can also be found from [8], [12].

5 Experiments

In this section, we empirically evaluate our algorithms in four real datasets and compare the performance with many existing algorithms according to some different criteria. The results show that our strategies are pretty well.

5.1 Datasets

In our experiments, we adopt the historical daily prices in stock markets, which can be easily obtained and hence available to other researchers. Data from other types of markets, such as high frequency intra-day quotes, currency and commodity markets, are either expensive or hard to obtain and process, which can reduce the experimental reproducibility. Table 3 summarizes the four real and diverse datasets from stock markets and index markets[5] employed in this paper.

The first dataset is the well-known NYSE dataset, one "standard" dataset pioneered by [11] and followed by most subsequent researchers on the field of online PS in [5], [7], [8], [9], [13], [21], [22], [23], [41]. This dataset contains 5,651 daily price relatives of 36 stocks in New York Stock Exchange (NYSE) for a 22-year period from Jul. 3rd 1962 to Dec. 31st 1984. We refer to it as "NYSE(O)".

5. All related codes and the datasets, including their compositions, are available on http://olps.stevenhoi.org/.

TABLE 3
Summary of Four Real Datasets

Data set	Region	Time Frame	#days	#assets
NYSE(O)	US	3/7/1962-31/12/1984	5651	36
NYSE(N)	US	1/1/1985-30/6/2010	6431	23
DJA	US	1/1/2001-14/1/2003	507	30
MSCI	Global	1/4/2006-31/3/2010	1043	24

The second dataset is the extended version of the above NYSE dataset and is collected by [23]. For consistency, this dataset is from Jan. 1st 1985 to Jun. 30th 2010, which consists of 6,431 trading days. We denote this dataset as "NYSE(N)" for short. It is worth noting that this new dataset consists of 23 stocks rather than the previous 36 stocks owing to amalgamations and bankruptcies.

The third dataset "DJA" is collected by [12], which consists of 30 stocks from Dow Jones Industrial Average containing price relatives of 507 trading days, ranging from Jan 1st 2011 to Jan 14th 2013.

The fourth dataset is "MSCI", a collection of global equity indices which are the constituents of MSCI World Index. It contains 24 indices which represent the equity markets of 24 countries around the world, and consists of a total of 1,043 trading days, ranging from Apr. 1st 2006 to Mar. 31st 2010.

The above testbed covers much long trading periods from 1962 to 2010 and diversified markets, which enables us to examine how the proposed RMR strategy performs under different events and crises. For example, it covers several well-known events in the stock markets, such as dot-com bubble from 1995 to 2000 and subprime mortgage crisis from 2007 to 2009. The first three datasets are chosen to test strategy capability on stocks, while the MSCI dataset aims to test the proposed strategy on global indices, which may be potentially applicable to "Fund of Funds" (FOF). As a remark, although we numerically test the RMR algorithm on stock markets, the proposed strategy could be generally applied to any type of financial markets.

5.2 Experimental Setup and Metrics

In this section, we give detailed experimental setup, including parameter seeting, performance measures and the transaction costs issue.

5.2.1 *Parameter Settings*

Regarding the parameter settings, there are two key parameters, i.e., w and ε in the proposed RMR algorithms and the variant version of RMR algorithm. Additionally, there is another key parameter c for the variant version of RMR algorithm. w represents the length of window, ε is about sensitivity parameter, and c is also the sensitivity parameter in the variant version. The parameter τ related with Preposition 2 is taken as 10.[6] Roughly speaking, the best values for these parameters are often dataset dependent. In the experiments, we simply set these parameters empirically without tuning for each dataset separately. Specifically, for all datasets and experiments, we set w to 5 and ε to 5 in the two algorithms and set c to 0.01 in the variant version. It is worth noting that these choices for parameters are not always the best. Our experiments on the parameter sensitivity in Section 5.4.4 show that the proposed algorithms are quite robust with respect to different parameter settings.

5.2.2 Performance Measures

One of the standard criteria to evaluate the performance of a strategy is *portfolio cumulative wealth* achieved by the strategy until the end of the whole trading period. In our study, we simply set the initial wealth $\mathbf{S_0} = 1$ and thus the notation $\mathbf{S_n}$ also denotes *portfolio cumulative wealth* at the end of the nth trading day, which is the ratio of the portfolio cumulative wealth divided by the initial wealth. Another equivalent criterion is *Annualized Percentage Yield* (APY) which takes the compounding effect into account, that is, $\mathrm{APY} = \sqrt[y]{S_n} - 1$, where y is the number of years corresponding to n trading days. *Winning Ratio* (WT) denotes the percentage of cases when the proposed strategy beats the BAH strategy. Typically, the higher the value of portfolio cumulative wealth or annualized percentage yield and WT, the more performance preferable the trading strategy is.

To test whether simple luck can generate the return of the proposed strategy, we can also conduct a statistical test to measure the probability of this situation, as is popularly done in the fund management industry [13], [44]. First, we separate the portfolio daily returns into two components: one benchmark-related and the other non-benchmark-related by regressing the portfolio excess returns against the benchmark excess returns. Formally, $s_t - s_t(F) = \alpha + \beta(s_t(B) - s_t(F)) + \epsilon(t)$, where s_t stands for the portfolio daily returns, $s_t(B)$ denotes the daily returns of the benchmark (market index) and $s_t(F)$ is the daily returns of the risk-free assets (here we simply choose Treasury bill and set it to 1.000156, or equivalently, annual interest of 4 percent). This regression estimates the portfolio's alpha(α), which indicates the performance of the investment after accounting for the involved risk. Then we conduct a statistical t-test to evaluate whether alpha is significantly different from zero, by using the t statistic $\frac{\alpha}{SE(\alpha)}$, where $SE(\alpha)$ is the standard error for the estimated alpha. Thus, by assuming the alpha is normally distributed, we can obtain the probability that the returns of the proposed strategy are generated by simple luck. Generally speaking, the smaller the probability, the higher confidence the trading strategy.

We also evaluate their performance by *risk* and *risk-adjusted return* of portfolios [45], [46]. One common way to achieve this is to use *annualized standard deviation* of daily returns to measure the volatility risk and *annualized Sharpe Ratio* (SR) to evaluate the risk-adjusted return. For risk-adjusted return, we calculate *annualized Sharpe Ratio* according to, $\mathrm{SR} = \frac{\mathrm{APY} - R_f}{\sigma_p}$, where R_f is the risk-free return (typically the return of Treasury bills, fixed at 4 percent in this work), and σ_p is the annualized standard deviation of daily returns. Basically, higher annualized Sharpe Ratios indicate better performance of a trading strategy concerning the volatility risk. We also adopt *Calmar Ratio* (CR) to measure the return relative of the drawdown risk of a portfolio, calculated as $\mathrm{CR} = \frac{\mathrm{APY}}{\mathrm{MDD}}$, where MDD is the *Maximum DrawDown* and measures the downside risk of different strategies. Generally speaking, higher Calmar Ratios indicate better performance of a trading strategy concerning the drawdown risk.

5.2.3 Practical Issue

In reality, an important and unavoidable issue is *transaction cost*. Generally, there are two ways to deal with this problem. The first is that the PS process does not consider the transaction cost while the following rebalancing incurs transaction costs and this method has been commonly adopted by learning to select portfolio strategies. The second way is that the transaction cost is directly involved in the PS process [31]. In our experiments, we take the first way and adopt *proportional transaction cost* model, which is proposed by [12], [30]. Specifically, rebalancing the portfolio incurs a transaction cost on every buy and sell operation with regarding to a transaction cost rate $\gamma \in (0, 1)$. At the beginning of the tth trading day, the portfolio manager rebalances the portfolio from the previous closing price adjusted portfolio $\hat{\mathbf{b}}_{t-1}$ to a new portfolio $\mathbf{b}_t$, incurring a transaction cost of $\frac{\gamma}{2} \times \sum_i |b_{(t,i)} - \hat{b}_{(t-1,i)}|$, where the initial portfolio is set to $(0, \ldots, 0)$. Thus, with transaction cost rate γ, the cumulative wealth achieved by the end of the nth trading day can be expressed as:

$$S_n = S_0 \prod_{t=1}^{n} [(\mathbf{b}_t^T \mathbf{x}_t) \times (1 - \frac{\gamma}{2} \times \sum_i |b_{(t,i)} - \hat{b}_{(t-1,i)}|)].$$

To the best of our knowledge, this model cannot work for high frequency data, since even a small rate will cause all methods approach zero.

5.3 Comparison Approaches

In our experiments, we implement the proposed RMR strategy and its variant, RMR-Variant. We compare them with a number of benchmarks and existing strategies as describe in Section 3. Below we summarize the list of compared algorithms.

1) Market: Market strategy that is uniform Buy-And-Hold strategy;
2) Best-Stock: Best stock in a market that is obviously a hindsight strategy;
3) BCRP: Best Constant Rebalanced Portfolios strategy in hindsight;
4) UP: Cover's Universal Portfolios implemented according to [47];
5) EG: Exponential Gradient with the best parameter $\eta = 0.05$ suggested by [9];
6) ONS: Online Newton Step with the parameters suggested by [7], that is, $\eta = 0, \beta = 1, \gamma = \frac{1}{8}$.
7) B^k: Nonparametric kernel-based moving window strategy with $W = 5, L = 10$ and threshold $c = 1.0$ for daily datasets that has the best empirical performance according to [21];
8) B^{NN}: Nonparametric nearest neighbor based strategy with parameter $W = 5, L = 10$, and $p_\ell = 0.02 + 0.5\frac{\ell - 1}{L - 1}$ as suggested by [22];
9) CORN: Correlation-driven nonparametric learning approach with parameter $W = 5$ and $\rho = 0.1$ suggested by [23];

TABLE 4
Cumulative Wealth Achieved by Various Strategies on the Four Datasets

Methods	NYSE(O)	NYSE(N)	DJA	MSCI
Market	14.50	18.06	0.76	0.91
Best-stock	54.14	83.51	1.19	1.50
BCRP	250.60	120.32	1.24	1.51
UP	26.68	31.49	0.81	0.92
EG	27.09	31.00	0.81	0.93
ONS	109.91	21.59	1.53	0.86
B^k	1.08E+09	4.64E+03	0.68	2.64
B^{NN}	3.35E+11	6.80E+04	0.88	13.47
CORN	1.48E+13	5.37E+05	0.84	**26.19**
Anticor	2.41E+08	6.21E+06	2.29	3.22
PAMR	5.14E+15	1.25E+06	0.68	15.23
CWMR	6.49E+15	1.41E+06	0.68	17.28
OLMAR	4.04E+16	2.24E+08	2.05	16.33
RMR	1.64E+17	$\mathbf{3.25E+08}$	**2.67**	16.76
RMR(max)	2.81E+17	4.73E+08	3.47	19.07
RMR-Var	$\mathbf{1.67E+17}$	3.24E+08	**2.67**	17.48
RMR-Var(max)	2.83E+17	4.75E+08	3.48	19.07

The best results (excluding the best experts at the bottom, which is in hindsight) in each dataset are highlighted in ***bold****.*

10) Anticor: BAH_{30} (Anticor) as a variant of Anticor to smooth the volatility, which is a better solution proposed by [12];
11) PAMR: Passive aggressive mean reversion strategy with parameter $\epsilon = 0.5$ suggested by [13];
12) CWMR: Confidence Weighted Mean Reversion Strategy with parameter $\phi = 2, \epsilon = 0.5$ suggested by [5];
13) OLMAR: Online Moving Average Reversion strategy with parameter $\epsilon = 5, w = 5$ [8];

5.4 Experimental Results

5.4.1 Experiment 1: Evaluation of Cumulative Wealth

We first compare performance of the competing approaches based on their cumulative wealth without considering the transaction cost and the results are illustrated in Table 4.

From the results, we can draw several observation. First of all, we find that almost all learning-to-trade algorithms can beat the market index, that is, the uniform BAH strategy, on all the datasets. This shows that it is promising to investigate learning-to-trade algorithms for PS. Second, the cumulative wealth achieved by RMR and RMR-Variant are similar since they both use the robust estimators. Third, compared with the existing mean reversion strategies (Anitor, PAMR, CWMR and OLMAR), RMR and RMR-Variant strategies obtained higher cumulative wealth on the datasets NYSE(O), NYSE(N) and DJA. Moreover, RMR-Variant get the better return than these mean reversion strategies on MSCI dataset.

Besides the preceding final cumulative wealth, we are also interested in examining how the total wealth achieved by various strategies change over different trading periods. In Fig. 2, we plot the wealth achieved by the proposed RMR algorithm, state-of-the-art algorithms (PAMR, CWMR, OLMAR), plus two benchmarks (Market and BCRP). As RMR-Variant perform similar to the RMR algorithm, we ignore it in these figures. From the results, we can see that the proposed RMR strategy consistently surpassed the benchmarks and the competing strategies over the entire trading period on most datasets, which again validates the efficacy of the proposed technique.

Finally, to measure whether the results are generated by simple luck, we conduct widely accepted statistical test as described in Section 5.2.2. Tables 5 and 6 respectively summarizes the statistical test results, which show that there almost exists no chance that the cumulative wealth is generated by luck. To be specific, the probabilities for achieving the excess returns by luck are almost 0. So, the results show that the RMR strategy is promising and reliable PS technique to achieve high return with high confidence. Besides, we can find that the *winning ratio* (WT) against Market strategy is bigger than 50 percent on the four daily datasets, which further shows the proposed strategies' advantages.

TABLE 5
Statistical Test of RMR

Stat.	NYSE(O)	NYSE(N)	DJA	MSCI
Size	5651	6431	507	1043
MER(RMR)	0.0077	0.0036	0.0024	0.0030
MER(Market)	0.0005	0.0005	-0.0004	0.0000
Winning Ratio	0.5721	0.5332	0.5503	0.5925
α	0.0071	0.0031	0.0030	0.0030
β	1.2718	1.1628	1.2427	1.1885
t-statistics	15.7325	7.4222	2.5217	5.8380
p-value	0.0000	0.0000	0.0060	0.0000

5.4.2 Experiment 2: APY, Volatility, Sharpe Ratio, MDD, and CR

We evaluated the performance of APY, volatility, annualized Sharpe Ratio, MDD and CR of the compared strategies

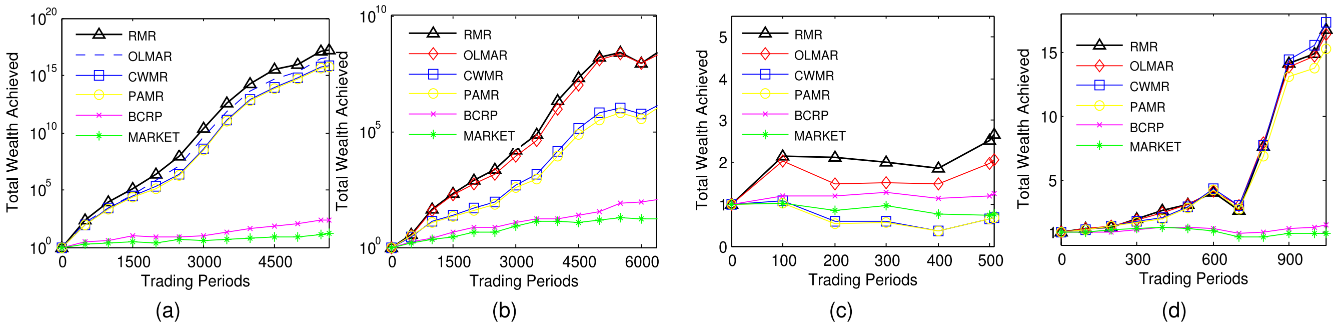

Fig. 2. Trend of cumulative wealth achieved by various strategies during the entire period on the four daily datasets (i.e., (a) NYSE(O), (b) NYSE(N), (c) DJA, and (d) MSCI).

TABLE 6
Statistical Test of RMR-Variant

Stat.	NYSE(O)	NYSE(N)	DJA	MSCI
Size	5651	6431	507	1043
MER(RMR-Var)	0.0077	0.0037	0.0025	0.0031
MER(Market)	0.0005	0.0005	-0.0004	0.0000
Winning Ratio	0.5721	0.5332	0.5503	0.5935
α	0.0070	0.0031	0.0030	0.0031
β	1.2602	1.1613	1.2325	1.1908
t-statistics	15.7949	7.5149	2.5906	5.9752
p-value	0.0000	0.0000	0.0049	0.0000

and summarize results in Table 7. From the results, we observe that on the NYSE(O), NYSE(N), and DJA datasets, both RMR and RMR-Variant algorithms achieved higher APYs and lower volatility than OLMAR. Moreover, the Sharpe Ratio of proposed strategies is also higher than that of OLMAR on the first three datasets. Although the volatility on dataset MSCI is higher for the proposed strategies, the APYs and Sharpe Ratios of the proposed strategies is higher than that of OLAMR. In addition, RMR and the variant version algorithms achieve lower MDD and higher CR on most datasets (except MSCI).

These encouraging results show that RMR is able to reach a good trade-off between return and risk, even though we do not explicitly consider risk in our formulation.

5.4.3 Experiment 3: Turnover

We use the turnover of the portfolio to measure the stability of the portfolio. Roughly speaking, turnover often measures what percentage of a portfolio's assets are bought and sold in a given year. Because the data in our four datasets spreads out many years, the turnover indicated here is the mean value of turnover of every trading period which is calculated by $\sum_{t=2}^{T} \frac{||b_t - \hat{b}_{t-1}||}{2(T-1)}$, where $||b_t - \hat{b}_{t-1}||/2$ is the turnover of one period. The portfolio manager rebalances the portfolio from the previous closing price adjusted portfolio $\hat{\mathbf{b}}_{t-1}$ to a new portfolio $\mathbf{b}_t$. In our experiment, we compare the turnover of RMR strategy with that of the state-of-the-art strategy (OLMAR). RMR and OLMAR are both multiple-period reversion strategies, so the comparison among them is more significant. Table 8 presents the explicit turnover of the portfolio of the strategies. As we observe, the turnover of the portfolio of RMR is smaller than that of OLMAR on the four datasets. Moreover, the RMR-Variant strategy achieve the similar turnover to RMR. Generally, the smaller turnover means that the portfolio is more stable, which can be attributed to the resistance to the noisy data or outliers. Thus, the small turnover empirically show the robustness of the proposed strategy. Furthermore, the smaller turnover usually results in less transaction cost. As analyzed in the above study, we get that RMR achieve higher wealth than OLMAR strategy when transaction cost is not considered. In the experiment, we know that the RMR achieve smaller turnover, thus, it may also achieve better results when transaction cost is taken into account.

TABLE 7
The Comparison of APY, Volatility, Sharpe Ratio, MDD, and CR Among OLMAR, RMR, and RMR-Variant Strategies

Criteria	Strategy	NYSE(O)	NYSE(N)	DJA	MSCI
	OLMAR	4.6862	1.0950	0.4301	1.0101
APY	RMR	5.0602	1.1250	0.6334	1.0234
	RMR-Var	5.0653	1.1249	0.6334	1.0449
	OLMAR	0.5646	0.5654	0.5225	0.3901
Volatility	RMR	0.5639	0.5644	0.5117	0.3913
	RMR-Var	0.5639	0.5644	0.5117	0.3929
	OLMAR	8.2295	1.8659	0.7467	2.4869
Sharpe	RMR	8.9031	1.9224	1.1597	2.5128
Ratio	RMR-Var	8.9112	1.9222	1.1597	2.5574
	OLMAR	0.4299	0.9329	0.4641	0.4552
MDD	RMR	0.4243	0.9096	0.3469	0.4933
	RMR-Var	0.4243	0.9098	0.3470	0.4933
	OLMAR	10.9012	1.1738	0.9268	2.2189
CR	RMR	11.9249	1.2368	1.8258	2.0747
	RMR-Var	11.9370	1.2365	1.8255	2.1183

TABLE 8
The Comparison of Turnover Among OLMAR, RMR, and RMR-Variant Strategies

Strategy	NYSE(O)	NYSE(N)	DJA	MSCI
OLMAR	72.7495%	68.4152%	70.6738%	73.8301%
RMR	68.1331%	63.4909%	65.0425%	69.3997%
RMR-Var	68.1325%	63.4898%	65.0451%	69.2816%

5.4.4 Experiment 4: Evaluation of Parameters Sensitivity

We now experimentally evaluate how different choices of parameters affect the cumulative wealth performance. RMR and RMR-Variant contain two parameters, that is, the sensitivity parameter ϵ and window size w. Additionally, another parameter c is needed for RMR-Variant strategy.

First, we examine the performance of the RMR algorithm by varying sensitivity parameter ϵ from 0 to 100 with fixed $w = 5$. Fig. 3 shows the effects of varied ϵ values for the RMR algorithm and two benchmarks Market and BCRP strategies on the four datasets. The results show that cumulative wealth sharply grows as ϵ increases and then flattens when ϵ crosses a threshold. Second, we evaluate the other important parameter for RMR algorithm, that is, window w. With fixed $\epsilon = 5$, Fig. 4 show the cumulative wealth of RMR algorithm by varying w from 3 to 100. The cumulative wealth decrease as w grows bigger on most datasets (except DJA). The two figure show that $\epsilon = 5$ and $w = 5$ are not the optimal parameters, and RMR is robust w.r.t. different parameters and it is convenient to choose satisfying parameters.

In addition, we evaluate another parameter c for RMR-Variant algorithm. As the sensitivity results of parameters ϵ and w for RMR-Variant are similar to the RMR algorithm, we only present sensitivity results of c for RMR-Variant algorithm. Fig. 5 shows the performance of the RMR-Variant algorithm by varying c from 0.001 to 98 with fixed $\epsilon = 5,\ w = 5$. The results show that the performance keeps stable when c is within the neighbor of two ends of c, which can be useful to select the optimal parameter.

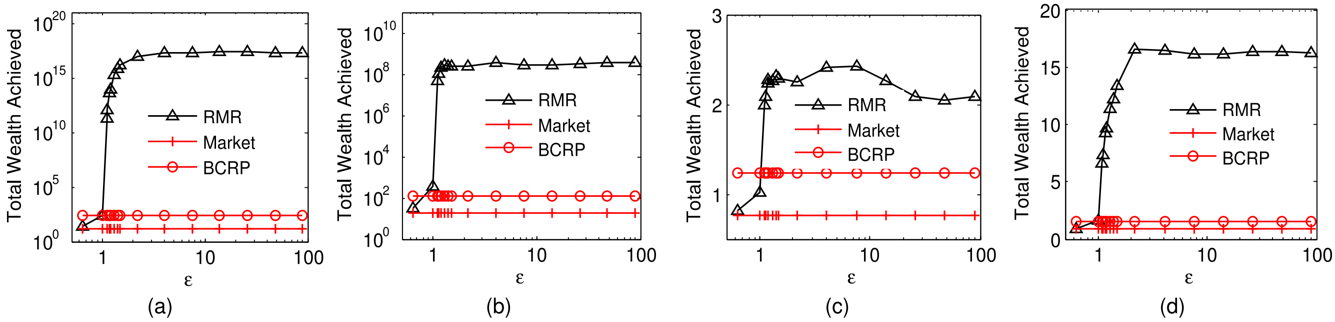

Fig. 3. Parameter sensitivity of RMR w.r.t. ϵ with fixed w (w=5) on the four datasets (i.e., (a) NYSE(O), (b) NYSE(N), (c) DJA, and (d) MSCI).

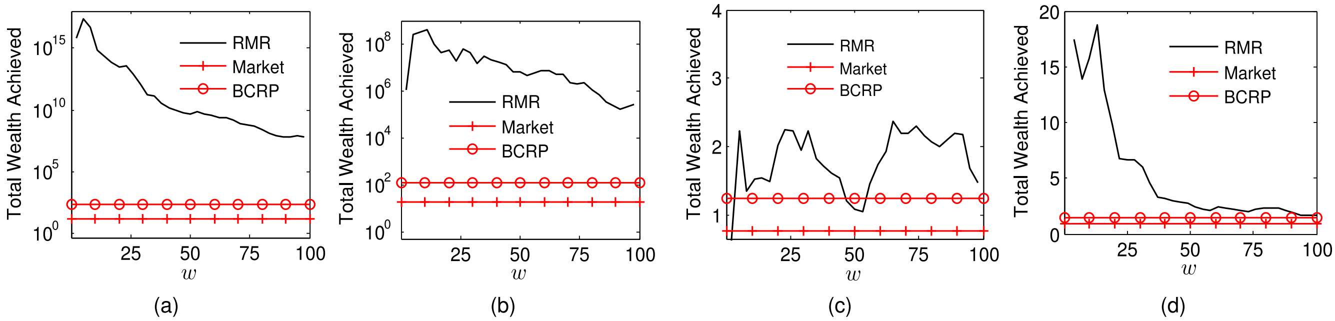

Fig. 4. Parameter sensitivity of RMR w.r.t. w with fixed ϵ ($\epsilon = 5$) on the four datasets (i.e., (a) NYSE(O), (b) NYSE(N), (c) DJA, and (d) MSCI).

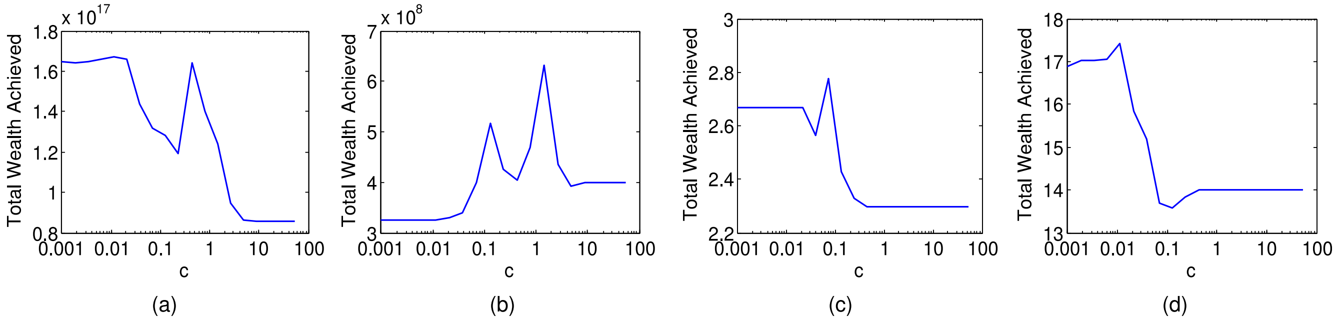

Fig. 5. Parameter sensitivity of RMR-Varint w.r.t. c with fixed $\epsilon,\ w$ ($\epsilon = 5$ and $w = 5$) on the four datasets (i.e., (a) NYSE(O), (b) NYSE(N), (c) DJA, and (d) MSCI).

5.4.5 *Experiment 5: Transaction Costs*

In practice, transaction cost is an important and unavoidable issue that should be addressed. In our experiment, we adopt *proportional transaction cost* model stated in Section 5.2.3. We test the effect of proportional transaction cost when the transaction cost rate γ varies from 0 to 1 percent, plus the cumulative wealth achieved by two benchmarks (Market and BCRP) and the state-of-the-arts (PAMR and OLMAR).

Figs. 6 and 7, present the results of RMR and RMR-Variant, respectively. As we can observe, the performance with transaction costs is market dependent. When the transaction cost increases, the total wealth achieved by RMR and RMR-Variant strategies drops considerably. However, we found that even with a rather high transaction cost, the two strategies still performs convincingly well. Compared with the benchmarks, the results clearly demonstrate that on all datasets, the two algorithms are fairly robust with respect to the transaction costs, where the break-even rates ranges from 0.3 to 0.9 percent. Thus, the proposed strategies can withstand moderate transaction costs even though we do not explicitly tackle it during the PS process.

6 Conclusion

In this paper, we propose a novel multiple period online PS strategy named "robust median reversion" , which exploits the reversion phenomenon of stock prices by robust L_1-median estimator and online learning technologies. The proposed approaches overcome the limitations of many existing online PS strategies that often suffer from noise and outliers in real-world markets. Our empirical studies show that the proposed RMR algorithm can substantially beat the market and the best stock, and also consistently suppasses a variety of state-of-the-art algorithms.

In the future, we plan to study the following aspects. Firstly, RMR's universality is still an open question, although this may not be required in real investment. Secondly, more financial issues need to be studied, for example, bankrup assets. It is interesting to study the behaviors of the bankrupt assets and design strategies to exploit them. Finally, though RMR handles the issue of transaction costs well, it is not formally addressed in our problem formulation. It would be interesting to incorporate the transaction cost issue when formulating the problem,

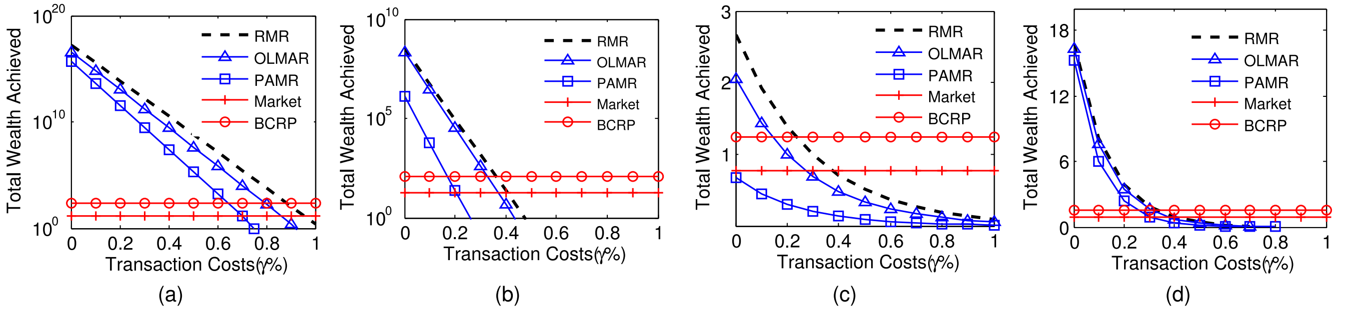

Fig. 6. Scalability of the total wealth achieved by RMR with respect to transaction cost rate γ percent on the four datasets (i.e., (a) NYSE(O), (b) NYSE(N), (c) DJA, and (d) MSCI).

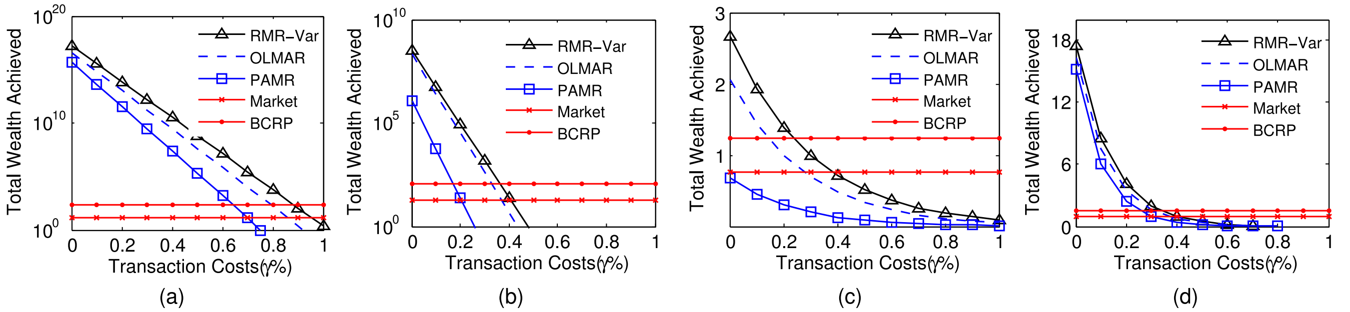

Fig. 7. Scalability of the total wealth achieved by RMR-Variant with respect to transaction cost rate γ percent on the four datasets (i.e., (a) NYSE(O), (b) NYSE(N), (c) DJA, and (d) MSCI).

especially, in case of high transaction costs and high frequency trading.

Appendix A
Proof of the Proposition 1

Lemma 1. *This lemma can be got from [17]. Let $\mathbf{x}_1, \ldots, \mathbf{x}_m$ be m distinct points in $\mathbb{R}^d$ and $\eta_1, \ldots, \eta_m$ be m positive numbers. Think of the η_is as weights of the $\mathbf{x}_i$s, and let $C(\mathbf{y})$ denote the weighted sum of distance of $\mathbf{y}$ from $\mathbf{x}_1, \ldots, \mathbf{x}_m$: $C(\mathbf{y}) = \sum_i \eta_i d_i(\mathbf{y})$, where $d_i(\mathbf{y}) = \|\mathbf{y} - \mathbf{x_i}\|$, the Euclidean distance between $\mathbf{y}$ and $\mathbf{x}_i$ in $\mathbb{R}^d$. Then a point $\mathbf{y} \in \mathbb{R}^d$ that minimizes the "cost function" $C(\mathbf{y})$, i.e., to find*

$$\begin{aligned}\mathbf{M} &= \mathbf{M}(\mathbf{x}_1, \ldots, \mathbf{x}_m; \eta_1, \ldots, \eta_m)\\ &= \arg\min\{C(\mathbf{y}) : \mathbf{y} \in \mathbb{R}^d\},\end{aligned} \tag{8}$$

can be calculated through iteration, and the iteration process is:

$$\mathbf{y} \to T(\mathbf{y}) = \left(1 - \frac{\eta(\mathbf{y})}{\gamma(\mathbf{y})}\right)^+ \tilde{T}(\mathbf{y}) + \min\left(1, \frac{\eta(\mathbf{y})}{\gamma(\mathbf{y})}\right)\mathbf{y},$$

with the convention 0/0=0 in the computation of $\eta(\mathbf{y})/\gamma(\mathbf{y})$, where $\tilde{T}(\mathbf{y})$ is as

$$\begin{aligned}\tilde{T}(\mathbf{y}) &= \left\{\sum_{\mathbf{x}_i \neq \mathbf{y}} \tfrac{\eta_i}{\|\mathbf{y}-\mathbf{x}_i\|}\right\}^{-1} \sum_{\mathbf{x}_i \neq \mathbf{y}} \tfrac{\eta_i \mathbf{x}_i}{\|\mathbf{y}-\mathbf{x}_i\|},\\ \eta(\mathbf{y}) &= \begin{cases}\eta_k & if\ \mathbf{y} = \mathbf{x}_k, \quad k = 1, \ldots, m\\ 0 & otherwise\end{cases},\\ r(\mathbf{y}) &= \|\tilde{R}(\mathbf{y})\|, \quad \tilde{R}(\mathbf{y}) = \sum_{\mathbf{x}_i \neq \mathbf{y}} \eta_i \tfrac{\mathbf{x}_i - \mathbf{y}}{\|\mathbf{x}_i - \mathbf{y}\|}.\end{aligned}$$

Proof. Based on Lemma 1, if we set $\eta_i = 1$, $m = k-1$, $\mathbf{x}_1 = \mathbf{p}_t, \ldots, \mathbf{x}_m = \mathbf{p}_{t-k+1}$, then Eq. (8) is same to the Eq. (2), and we can get the conclusion of proposition 1:

$$\begin{aligned}\tilde{T}(\mathbf{y}) &= \left\{\sum_{\mathbf{p}_{t-i} \neq \mathbf{y}} \tfrac{1}{\|\mathbf{y}-\mathbf{p}_{t-i}\|}\right\}^{-1} \sum_{\mathbf{x}_i \neq \mathbf{y}} \tfrac{\mathbf{p}_{t-i}}{\|\mathbf{y}-\mathbf{p}_{t-i}\|},\\ \eta(\mathbf{y}) &= \begin{cases}1 & if\ \mathbf{y} = \mathbf{p}_{t-i}, \quad i = 0, \ldots, k-1\\ 0 & otherwise\end{cases},\\ \tilde{R}(\mathbf{y}) &= \sum_{\mathbf{p}_{t-i} \neq \mathbf{y}} \tfrac{\mathbf{p}_{t-i} - \mathbf{y}}{\|\mathbf{p}_{t-i} - \mathbf{y}\|}.\end{aligned}$$

□

Appendix B
Proof of the Proposition 3

Proof. If $\hat{\mathbf{x}}_{t+1}^T \mathbf{b}_t - \varepsilon \geq 0$, then $\mathbf{b} = \mathbf{b}_t$. If $\hat{\mathbf{x}}_{t+1}^T \mathbf{b}_t - \varepsilon < 0$, then

$$\begin{aligned}L(\mathbf{b}, \alpha, \lambda) = \ & \tfrac{1}{2}\|\mathbf{b} - \mathbf{b}_t\|^2 + \alpha(\hat{\mathbf{x}}_{t+1}^T \mathbf{b}_t - \varepsilon)\\ & + \lambda(\mathbf{b}^T \mathbf{1} - 1)\end{aligned} \tag{9}$$

so,

$$\frac{\partial L}{\partial \mathbf{b}} = (\mathbf{b} - \mathbf{b}_t) + \alpha \hat{\mathbf{x}}_{t+1} + \lambda \mathbf{1} = 0 \tag{10}$$

$$\lambda = -\frac{\alpha \mathbf{1}^T \hat{\mathbf{x}}_{t+1}}{d} = -\alpha \overline{x}_{t+1} \mathbf{1}. \tag{11}$$

Substituting Eqs. (11) into (10) leads to

$$\mathbf{b} = \mathbf{b}_t - \alpha(\hat{\mathbf{x}}_{t+1} - \overline{x}_{t+1}\mathbf{1}). \tag{12}$$

To substitute Eqs. (11) and (12) into Eq. (9), we can get

$$\begin{aligned}L =& \frac{1}{2}\alpha^2 \|\hat{\mathbf{x}}_{t+1} - \overline{x}_{t+1}\mathbf{1}\|^2 + \alpha[\hat{\mathbf{x}}_{t+1}^T(\mathbf{b}_t - \alpha(\hat{\mathbf{x}}_{t+1}\\ & - \overline{x}_{t+1}\mathbf{1})) - \varepsilon]\\ =& \frac{1}{2}\alpha^2 \|\hat{\mathbf{x}}_{t+1} - \overline{x}_{t+1}\mathbf{1}\|^2 + \alpha \hat{\mathbf{x}}_{t+1}^T \mathbf{b}_t - \alpha^2(\hat{\mathbf{x}}_{t+1}^T \hat{\mathbf{x}}_{t+1}\\ & - \hat{\mathbf{x}}_{t+1}^T \overline{x}_{t+1}\mathbf{1}) - \alpha\varepsilon.\end{aligned}$$

Moreover,

$$
\begin{aligned}
&\|\hat{\mathbf{x}}_{t+1}-\overline{x}_{t+1}\mathbf{1}\|^2\\
=&(\hat{\mathbf{x}}_{t+1}-\overline{x}_{t+1}\mathbf{1})^T(\hat{\mathbf{x}}_{t+1}-\overline{x}_{t+1}\mathbf{1})\\
=&(\hat{\mathbf{x}}_{t+1}^T-\overline{x}_{t+1}\mathbf{1}^T)(\hat{\mathbf{x}}_{t+1}-\overline{x}_{t+1}\mathbf{1})\\
=&\hat{\mathbf{x}}_{t+1}^T\hat{\mathbf{x}}_{t+1}-\hat{\mathbf{x}}_{t+1}^T\overline{x}_{t+1}\mathbf{1}-\overline{x}_{t+1}\mathbf{1}^T\hat{\mathbf{x}}_{t+1}+\overline{x}_{t+1}\mathbf{1}^T\overline{x}_{t+1}\mathbf{1}\\
=&\hat{\mathbf{x}}_{t+1}^T\hat{\mathbf{x}}_{t+1}-\hat{\mathbf{x}}_{t+1}^T\overline{x}_{t+1}\mathbf{1}-\overline{x}_{t+1}\mathbf{1}^T\hat{\mathbf{x}}_{t+1}+\overline{x}_{t+1}^2 d\\
=&\hat{\mathbf{x}}_{t+1}^T\hat{\mathbf{x}}_{t+1}-\hat{\mathbf{x}}_{t+1}^T\overline{x}_{t+1}\mathbf{1}-\overline{x}_{t+1}\mathbf{1}^T\hat{\mathbf{x}}_{t+1}+\overline{x}_{t+1}\mathbf{1}^T\hat{\mathbf{x}}_{t+1}\\
=&\hat{\mathbf{x}}_{t+1}^T\hat{\mathbf{x}}_{t+1}-\hat{\mathbf{x}}_{t+1}^T\overline{x}_{t+1}\mathbf{1}.
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
L =&\frac{1}{2}\alpha^2\|\hat{\mathbf{x}}_{t+1}-\overline{x}_{t+1}\mathbf{1}\|^2+\alpha\hat{\mathbf{x}}_{t+1}^T\mathbf{b}_t\\
&-\alpha^2\|\hat{\mathbf{x}}_{t+1}-\overline{x}_{t+1}\mathbf{1}\|^2-\alpha\varepsilon\\
=&-\frac{1}{2}\alpha^2\|\hat{\mathbf{x}}_{t+1}-\overline{x}_{t+1}\mathbf{1}\|^2+\alpha\hat{\mathbf{x}}_{t+1}^T\mathbf{b}_t-\alpha\varepsilon\\
\frac{\partial L}{\partial\alpha}=&-\alpha\|\hat{\mathbf{x}}_{t+1}-\overline{x}_{t+1}\mathbf{1}\|^2+\hat{\mathbf{x}}_{t+1}^T\mathbf{b}_t-\varepsilon=0\\
\alpha =&\frac{\hat{\mathbf{x}}_{t+1}^T\mathbf{b}_t-\varepsilon}{\|\hat{\mathbf{x}}_{t+1}-\overline{x}_{t+1}\mathbf{1}\|^2}.
\end{aligned}
$$

Therefore, we have the result of Proposition 3:

$$\mathbf{b}_{t+1}=\mathbf{b}_t-\alpha_{t+1}(\hat{\mathbf{x}}_{t+1}-\overline{x}_{t+1}\cdot\mathbf{1}),$$

where α_{t+1} is the Lagrangian multiplier calculated as

$$\alpha_{t+1}=\min\left\{0,\frac{\hat{\mathbf{x}}_{t+1}^T\mathbf{b}_t-\epsilon}{\|\hat{\mathbf{x}}_{t+1}-\overline{x}_{t+1}\cdot\mathbf{1}\|^2}\right\}.$$

□

ACKNOWLEDGMENTS

This work was partially done when the first author was visiting the Computer Science Department, University of California Santa Cruz, he would like to thank Professor Manfred Warmuth for his warm invitation and hospitality. The work was partially supported by the National Natural Science Foundation of China (11501204, 71401128), the Natural Science Foundation of Shanghai (15ZR1408300), the Key Projects of Fundamental Research Program of Shanghai Municipal Commission of Science and Technology (14JC1400300), Shanghai Key Laboratory of Intelligent Information Processing (IIPL-2014-001), the special Postdoctoral Science Foundation of China (201104247), the project of SRF for ROCS, SEM and Singapore Ministry of Education Academic Research Fund Tier 1 Grant (14-C220-SMU-016).

REFERENCES

[1] D. Huang, J. Zhou, B. Li, S. Hoi, and S. Zhou, "Robust median reversion strategy for on-line portfolio selection," in *Proc. Int. Joint Conf. Artif. Intell.*, 2013, pp. 2006–2012.

[2] H. Markowitz, "Mean-variance analysis in portfolio choice and capital markets," *J. Finance*, vol. 7, no. 1, pp. 77–91, 1952.

[3] J. J. Kelly, "A new interpretation of information rate," *AT&T Tech. J.*, vol. 35, pp. 917–926, 1956.

[4] P. Das and A. Banerjee, "Meta optimization and its application to portfolio selection," in *Proc. Int. Conf. Knowl. Discovery Data Mining*, 2011, pp. 1163-1171.

[5] B. Li, S. C. H. Hoi, P. L. Zhao, and V. Gopalkrishnan, "Confidence weighted mean reversion strategy for online portfolio selection," *ACM Trans. Knowl. Discovery Data*, vol. 7, no. 1, pp. 4:1–4:38, 2013.

[6] B. Li and S. Hoi, "On-line portfolio selection: A survey," *ACM Comput. Surv.*, vol. 36, pp. 35:1–35, 2014.

[7] A. Agarwal, E. Hazan, S. Kale, and R. Schapire, "Algorithms for portfolio management based on the newton method," in *Proc. Int. Conf. Mach. Learning*, 2006, pp. 9–16.

[8] B. Li and S. C. H. Ho, "On-line portfolio selection with moving average reversion," in *Proc. Int. Conf. Mach. Learning*, 2012, pp. 273–228.

[9] D. Helmbold, R. Schapire, Y. Singer, and M. Warmuth, "On-line portfolios portfolio selection using multiplicative updates," *Math. Finance*, vol. 8, no. 4, pp. 325–347, 1998.

[10] N. Jegadeesh, "Evidence of predictable behavior of security returns," *J. Finance*, vol. 45, no. 3, pp. 881–898, 1990.

[11] T. Cover, "Universal portfolios," *Math. Finance*, vol. 1, pp. 1–29, 1991.

[12] A. Borodin, R. El-Yaniv, and V. Gogan, "Can we learn to beat the best stock," *JAIR*, vol. 21, pp. 579–594, 2004.

[13] B. Li, P. Zhao, S. C. Hoi, and V. Gopalkrishnan, "PAMR: Passive aggressive mean reversion strategy for portfolio selection," *Mach. Learn.*, vol. 87, no. 2, pp. 221–258, May 2012.

[14] R. C. Merton, "On estimating the expected return on a market: An exploratory investigation," *J. Financial Econ.*, vol. 8, pp. 323–361, 1980.

[15] A. Weber, *Uber den Standort der Industrien*. Tubingen : Mohr, 1909.

[16] E. Weiszfeld, "Sur le point pour lequel la somme des distances de n points donnes est minimum," *Tohoku Math. J.*, vol. 43, pp. 355–386, 1937.

[17] Y. Vardi and C. H. Zhang, "The multivariate l_1-median and associated data depth," *PNAS*, vol. 97, no. 4, pp. 1423–1426, 2000.

[18] B. Li and S. C. H. Hoi, *Online Portfolio Selection: Principles and Algorithms*. Boca Raton, FL, USA: CRC Press, 2015.

[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.

[20] A. A. Gaivoronski and F. Stella, "Stochastic nonstationary optimization for finding universal portfolios," *Ann. Operations Res.*, vol. 100, pp. 165–188, 2000.

[21] L. Györfi, G. Lugosi, and F. Udina, "Nonparametric kernel-based sequential investment strategies," *Math. Finance*, vol. 16, no. 2, pp. 337–357, 2006.

[22] L. Györfi, F. Udina, and H. Walk, "Nonparametric nearest neighbor based empirical portfolio selection strategies," *Statist. Decisions*, vol. 26, no. 2, pp. 145–157, 2008.

[23] B. Li, S. C. H. Hoi, and V. Gopalkrishnan, "Corn: Correlation-driven nonparametric learning approach for portfolio selection," *ACM Trans. Intelligent Syst. and Technol.*, vol. 2, no. 3, pp. 21:1–21:29, 2011.

[24] L. Györfi, G. Ottucsk, and H. Walk, Eds., *Machine Learning for Financial Engineering Advances in Computer Science and Engineering*. London, U.K.: Imperial College Press, 2012.

[25] P. Das, N. Johnson, and A. Banerjee, "Online lazy updates for portfolio selection with transaction costs," in *Proc. Nat. Conf. Artif. Intell.*, 2013, pp. 202–208.

[26] P. Das, N. Johnson, and A. Banerjee, "Online portfolio selection with group sparsity," in *Proc. 28th Nat. Conf. Artif. Intell.*, 2014, pp. 1185–1191.

[27] S. Kozat and A. Singer, "Universal semiconstant rebalanced portfolios," *Math. Finance*, vol. 21, no. 2, pp. 293–311, 2011.

[28] D. Huang, Y. Zhu, B. Li, S. Zhou, and S. Hoi, "Semi-universal portfolios with transaction costs," in *Proc. Int. Joint Conf. Artif. Intell.*, 2015, pp. 178–184.

[29] E. O. Thorp, "Portfolio choice and the kelly criterion," *Business and Economics Section of the American Statistical Association*, 1971.

[30] A. Blum and A. Kalai, "Universal portfolios with and without transaction costs," *Mach. Learning*, vol. 35, no. 3, pp. 193–205, 1999.

[31] L. Györfi and I. Vajda, "Growth optimal investment with transaction costs," in *Proc. Int. Conf. Algorithmic Learning Theory.*, 2008, pp. 108–122.

[32] E. Fama, "The behavior of stock market prices," *Finance Anal. J.*, vol. 51, no. 1, pp. 55–59, 1965.

[33] K. Grammer, O. Dekel, J. Keshet, S. Shalev-Shwartz, and Y. Singer, "Online passive-aggressive algorithms," *J. Mach. Learning Res.*, vol. 7, pp. 551–585, 2006.

[34] B. Brown, "Statistical use of spatial median," *J. R. Stat. Soc. B*, vol. 45, pp. 25–30, 1983.

[35] C. G. Small, "A survey of multidimensional medians," *Int. Stat. Rev.*, vol. 58, no. 3, pp. 263–277, 1990.
[36] V. Demiguel and F. Nogales, "Portfolio selection with robust estimation," *Operations Res.*, vol. 57, pp. 560–277, 2009.
[37] P. J. Huber, "Robust estimation of a location parameter," *Ann. Statist.*, vol. 53, pp. 73–101, 1964.
[38] H. Lopuhaa and P. Rousseeuw, "Breakdown points of affine equivariant estimators of multivariate location and covariance matrices," *Ann. Statist.*, vol. 19, no. 1, pp. 229–248, 1991.
[39] W. C. Davidon, "Variable metric method for minimization," *SIAM J. Optimization*, vol. 1, no. 1, pp. 1–17, 1991.
[40] C. G. Broyden, "The convergence of a class of double-rank minimization algorithms," *J. Inst. Math. Appl.*, vol. 6, pp. 76–90, 1970.
[41] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.
[42] J. Duchi, S. Shalev-Shwartz, Y. Singer, and T. Chandra, "Efficient projectinos onto the l1-ball for learning in high dimensions," in *Proc. Int. Conf. Mach. Learning*, 2008, pp. 272–279.
[43] I. Aldridge, *High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems*. Hoboken, NJ, USA: Wiley, 2010.
[44] R. Grinold and R. Kahn, Eds., *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk*. New York, NY, USA: McGraw-Hill, 1999.
[45] W. F. Sharpe, "A simplified model for portfolio analysis," *Manage. Sci.*, vol. 9, pp. 277–293, 1963.
[46] W. F. Sharpe, "The sharpe ratio," *Portfolio Manage.*, vol. 21, no. 1, pp. 49–58, 1964.
[47] A. Kalai and S. Vempala, "Efficient algorithms for universal portfolios," *J. Mach. Learning Res.*, vol. 3, pp. 423–440, 2002.